

# Antibody nomenclature

## From IMGT-ONTOLOGY to INN definition

Marie-Paule Lefranc

IMGT®, the international ImMunoGeneTics information system®; Université Montpellier and CNRS; Laboratoire d'ImmunoGénétique Moléculaire LIGM; Institut de Génétique Humaine IGH; Montpellier, France

An ontology is a formal representation of a knowledge domain. IMGT-ONTOLOGY was created to manage the complexity of immunogenetics knowledge and was the first ontology in the domain.<sup>1</sup> IMGT-ONTOLOGY has been essential for the creation in 1989 of IMGT®, the international ImMunoGeneTics information system® ([www.imgt.org](http://www.imgt.org)) and is now acknowledged as the global reference in immunogenetics and immunoinformatics.<sup>2</sup> IMGT-ONTOLOGY manages the immunogenetics knowledge through diverse facets that rely on the seven axioms of the Formal IMGT-ONTOLOGY or IMGT-Kaleidoscope: IDENTIFICATION, DESCRIPTION, CLASSIFICATION, NUMEROTATION, LOCALIZATION, ORIENTATION and OBTENTION.<sup>3,4</sup> These axioms postulate that objects, processes and relations have to be identified, described, classified, numerotated, localized and oriented and that the way they are obtained has to be determined. Key concepts of classification and numerotation necessary for antibody nomenclature will be briefly presented.

The genes that code the immunoglobulins (IG) or antibodies and T-cell receptors (TR) belong to highly polymorphic multigenic families. A major contribution of IMGT-ONTOLOGY was to set the principles of their classification and to propose a standardized nomenclature based on the concepts of “group,” “subgroup,” “gene” and “allele.”<sup>5,6</sup> The “group” concept classifies, for the IG and TR, a set of genes of the same “genotype:” variable (V), diversity (D), joining (J) or constant (C). The “subgroup” concept classifies a subset of genes that belong to the same group, and which, in a given species, share at least 75% of identity at the nucleotide sequence level (and in the germline configuration for the

V, D and J genes). The “gene” concept classifies a unit of DNA sequence that can be potentially transcribed and/or translated (this definition includes the regulatory elements in 5' and 3' and the introns, if present). In IMGT-ONTOLOGY, a gene name is composed of the name of the species and of the gene symbol, for example, *Homo sapiens* IGLV1-2. The “allele” concept classifies a polymorphic variant of a gene. Alleles identified by the mutations of the nucleotide sequence are classified by reference to allele \*01. They are reported in Alignments of alleles in IMGT Repertoire ([www.imgt.org](http://www.imgt.org)) and in IMGT/GENE-DB.<sup>7</sup>

The IMGT® gene nomenclature for IG and TR genes was approved at the international level by the Human Genome Organisation (HUGO) Nomenclature Committee (HGNC) in 1999 and endorsed by the World Health Organization-International Union of Immunological Societies (WHO-IUIS).<sup>8,9</sup> The IMGT® IG and TR gene names are the official reference for the genome projects and, as such, have been entered in IMGT/GENE-DB, in the Human Genome Database (GDB), in LocusLink at the National Center for Biotechnology Information (NCBI), in Entrez Gene when this database superseded LocusLink, in Ensembl at the European Bioinformatics Institute (EBI), and in the Vertebrate Genome Annotation (Vega) Browser at the Wellcome Trust Sanger Institute.

The IMGT-ONTOLOGY concepts of numerotation determine the principles of a unique numbering for a domain (sequences and 3D structures). The “IMGT unique numbering” is crucial for the antibody V domain and C domain standardized characterization.<sup>10,11</sup>

The “IMGT unique numbering for V domain” determines the “CDR-IMGT length” and “FR-IMGT length” (shown in IMGT® tools and databases between brackets and separated by dots).<sup>10</sup> The “IMGT unique numbering for C domain” determines the “loop length” and the “strand length.”<sup>11</sup> The “IMGT unique numbering” is illustrated by the “IMGT Collier de Perles,” which is a graphical representation in two dimensions (2D) of the amino acid sequences of the antibody V and C domains.<sup>12,13</sup> This concept is largely recognized at the international level and the expression “IMGT Collier de Perles” is now used in scientific publications.

Antibodies represent a large number of the pharmaceutical substances submitted to the WHO/International Nonproprietary Names (WHO/INN) Programme.<sup>14</sup> The INN definition of antibodies is based on the IMGT-ONTOLOGY concepts of classification (nomenclature), description (labels) and numerotation (IMGT unique numbering). Thus the INN definition provides the closest V and J genes and alleles corresponding to the amino acid sequences with, for the humanized (-zumab INN suffix) and human (-umab INN suffix) antibodies, the percentage of identity of the V regions. Amino acid changes of the constant region, by comparison with the reference alleles (and expected allotypes), and those of the FR4-IMGT, by comparison with the closest germline J genes and alleles, are indicated based on the IMGT numbering per domain.<sup>15</sup> This information is obtained with the IMGT/DomainGapAlign tool ([www.imgt.org](http://www.imgt.org)), which aligns the amino sequences with the IMGT domain reference directory.<sup>16</sup> Results from this tool are widely used for

antibody engineering and humanization design as they allow to precisely define and to easily compare amino acid sequences of the FR-IMGT and CDR-IMGT, between the nonhuman (mouse, rat, etc.) and the closest human V domains.

Since 2008, amino acid sequences of monoclonal antibodies (mAb, suffix -mab) and of fusion proteins for immune applications (FPIA, suffix -cept) from WHO/INN have been entered in IMGT®. These therapeutic applications emphasize the importance of the IMGT-ONTOLOGY concepts in bridging the gap between antibody sequences, 2D and 3D structures.

## References

1. Giudicelli V, Lefranc MP. Ontology for Immunogenetics: IMGT-ONTOLOGY, *Bioinformatics* 1999; 15:1047-54.
2. Lefranc MP, Giudicelli V, Ginestoux C, Jabado-Michaloud J, Folch G, Bellahcene F, et al. IMGT®, the international ImMunoGeneTics information system®. *Nucl Acids Res* 2009; 37:1006-12.
3. Duroux P, Kaas Q, Brochet X, Lane J, Ginestoux C, Lefranc MP, Giudicelli V. IMGT-Kaleidoscope, the Formal IMGT-ONTOLOGY paradigm. *Biochimie* 2008; 90:570-83.
4. Lefranc MP, Giudicelli V, Regnier L, Duroux P. IMGT, a system and an ontology that bridge biological and computational spheres in bioinformatics. *Brief Bioinform* 2008; 9:263-75.
5. Lefranc MP, Lefranc G. The Immunoglobulin FactsBook. Academic Press, London UK 2001; 1-458.
6. Lefranc MP, Lefranc G. The T cell receptor FactsBook. Academic Press, London UK 2001; 1-398.
7. Giudicelli V, Chaume D, Lefranc MP. IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucl Acids Res* 2005; 33:256-61.
8. Lefranc MP. WHO-IUIS Nomenclature Subcommittee for immunoglobulins and T cell receptors report. *Immunogenetics* 2007; 59:899-902.
9. Lefranc MP. WHO-IUIS Nomenclature Subcommittee for immunoglobulins and T cell receptors report August 2007 13<sup>th</sup> International Congress of Immunology, Rio de Janeiro, Brazil. *Dev Comp Immunol* 2008; 32:461-3.
10. Lefranc MP, Pommié C, Ruiz M, Giudicelli V, Foulquier E, Truong L, et al. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Dev Comp Immunol* 2003; 27:55-77.
11. Lefranc MP, Pommié C, Kaas Q, Duprat E, Bosc N, Guiraudou D, et al. IMGT unique numbering for immunoglobulin and T cell receptor constant domains and Ig superfamily C-like domains. *Dev Comp Immunol* 2005; 29:185-203.
12. Ruiz M, Lefranc MP. IMGT gene identification and Colliers de Perles of human immunoglobulin with known 3D structures. *Immunogenetics* 2002; 53:857-83.
13. Kaas Q, Ehrenmann F, Lefranc MP. IG, TR, MHC, IgSF and MhcSF: What do we learn from the IMGT Colliers de Perles? *Brief Funct Genomic Proteomic* 2007; 6:253-64.
14. General policies for monoclonal antibodies World Health Organization. INN Working Document 09.251. Update 18/12/2009. <http://www.who.int/medicines/services/inn/en/>
15. Jefferis R, Lefranc MP. Human immunoglobulin allotypes: Possible implications for immunogenicity. *Mabs* 2009; 1:332-8.
16. Ehrenmann F, Kaas Q, Lefranc MP. IMGT/3Dstructure-DB and IMGT/DomainGapAlign: a database and a tool for immunoglobulins or antibodies, T cell receptors, MHC, IgSF and MhcSF. *Nucl Acids Res* 2010; 38:301-7.